Novel DNA-based Vaccine Against the Encephalitis Alphaviruses

Field of the Invention

This invention relates to the cloning, sequencing and expression of the structural genes of western equine encephalitis (WEE) virus strain 71V-1658 and the development and use of the DNA-based vaccine against WEE.

Background of the Invention

List of Prior Art Literatures

Ausubel, F.M., et al, editors. (1995). *Current Protocols in Molecular Biology*, New York: John Wiley & Sons.

Bell, J. R., Bond, M.W., Nukapiller, M. B., Strauss, E.G., Strauss, J. H., Yamamoto, K., & Simizu, B. (1983). Structural proteins of western equine encephalitis virus: amino acid compositions and N-terminal sequences. *Journal of Virology* **45**, 708-714.

Bird, B.R. & Forrester, F.T. (1981). *Basic Laboratory Techniques In Cell Culture.* Atlanta: U. S. Department of Health and Human Services, Centers for Disease Control.

Calisher, C.H. & Karabatsos, N. (1988). Arbovirus serogroups: definition and geographic distribution. In *The Arboviruses: Epidemiology and Ecology*, Vol. I,.pp. 19-57. Edited by T. P Monath. CRC Press: Boca Raton, Fl.

Calisher, C. H., Shope, R.E, Brandt, W., Casals, J., Karabatsos, N., Murphy, F.A., Tesh, R.B., & Wiebe,,M.E.. (1980). Proposed antigenic classification of registered arbovirusess. *Intervirology* **14**, 229-232.

Calisher, C. H., Karabatsos, N., Lazuick, J. S.. Monath, T.P., & Wolff, K.L. (1988).

Reevaluation of the western equine encephalitis antigenic complex of alphaviruses (family Togaviridae) as determined by neutralization tests. *American Journal of Tropical Medicine and Hygiene* **38**, 447-452.

Cilnis, M.J., Kang, W. & Weaver, S.C. (1996). Genetic conservation of Highlands J viruses. *Virology* **218**, 343-351.

Frohman, M.A., Dush, M.K. & Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Science USA* **85**, 8998-9002.

Hahn, C. S., Lustig, S., Strauss, E.G. & Strauss, J.H. (1988). Western Equine Encephalitis virus is a recombinant virus. *Proceedings of the National Academy of Science USA* **85**, 5997-6001.

Johnson, R.E. & Peters, C.J. (1996). Alphaviruses. In *Fields Virology*, 3rd edn, pp. 843-898. Edited by B. N. Fields, *et al.*, New York: Raven Press.

Kuhn, R., Hong, Z. & Strauss, J.H. (1990). Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. *Journal of Virology* **64**, 1465-1476.

Kuhn, R.J., Niesters, H.G.M., Hong, Z. & Strauss, J.H. (1991). Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis. *Virology* **182**, 430-441.

Krieg, A.M., Yi, A.-K., Schorr, J. and Davis, H.L. (1998). The role of CpG dinucleotides in DNA vaccines. *Trends Microbiol.* **6**, 23-27.

McCluskie, M.J., Davies, H.L. (1999). Novel strategies using DNA for the induction of mucosal immunity. *Critic. Rev. in Immunol.* **19**, 303-329.

Ou, J.-H., Trent, D.W. & Strauss, J.H. (1982). The 3' non-coding regions of alphavirus RNAs contain repeating sequences. Journal of Molecular Biology **156**, 719-730.

Ou, J-H., Strauss, E. G. & Strauss, J.H. (1983). The 5' terminal sequences of the genomic RNAs of several alphaviruses. Journal of Molecular Biology **168**, 1-15.

Pardoll, DR, Beckering, AM. (1997). Exposing the immunology of naked DNA vaccines. *Immunity* **3**;165-169.

Pfeffer, M., Proebster, B., Kinney, R.M. & Kaaden, O-R. (1997). Genus-specific detection of alphaviruses by a semi-nested reverse transcription reaction. American Journal of tropical Medicine and Hygiene **57**, 709-718.

Pfeffer, M., Kinney, R.M. & Kaaden, O-R. (1998). The alphavirus 3'-nontranslated region: Size heterogeneity and arrangement of repeated sequence elements. Virology **240**, 100-108.

Prayaga, S.K., Fuller, D.H., Haynes, J.R. & Murphey-Corb, M. (1995). Particle-mediated nucleic acid immunization. *Vaccines* **95**, 105-109.

Reisen, W.K. & Monath, T.P. (1988). Western equine encephalomyelitis, pp. 89-137. *In The Arboviruses: Epidemiology and Ecology, Vol. V.* Edited by T.P. Monath. CRC Press: Boca Raton, Fl.

Robinson, H.L., Feltquate, D.M., Morin, M.J., Haynes, J.R., Webster, R.G. (1995). DNA vaccines: A new approach to immunization. *Vaccine* **95**:69-75.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular Cloning, a Laboratory Manual, 2nd edn..* Cold Spring Harbor: Cold Spring Harbor Laboratory.

Schlesinger, S. & Schlesinger, M.J. (1996). Togaviridae: The viruses and their replication, In *Fields Virology,* 3rd edn, pp. 825-841. Edited by B. N. Fields, *et al.*. New York: Raven Press.

Strauss, J. H., & Strauss, E.G. (1988). Evolution of RNA viruses. *Annual Review of Microbiology* **42**, 657-683.

Strauss, J. H., & Strauss, E.G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiological Review* **58**, 491-562.

Strauss, E.G., Rice, C.M. & Strauss, J.H. (1983). Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. *Proceedings of the National Academy of Science USA* **80**, 5271-5275.

Strauss, E.G., Rice, C.M. & Strauss, J.H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133**, 92-110.

Trent, D.W., & Grant, J.A. (1980). A comparison of new world alphaviruses in the western equine encephalomyelitis complex by immunochemical and oligonucleotide fingerprint techniques. *Journal or General Virology* **47**:261-282.

Weaver, S. C., Hagenbaugh, A., Bellew, L.A., Netesov, S. V. , Volchokov, V. I. , Chang, G.-J J., Clarke, D. K., Gousset, L., Scott, T. W., Trent, D. W. & Holland, J. J. (1993). A comparison of the nucleotide seqeunces of eastern and western equine encephalomyelitis viruses with those of other alphaviruses and related RNA viruses. *Virology* **197**, 375-390.

Weaver, S. C., Kang, W, Shirako, Y., Rumenapf, T., Strauss, E.G. & Strauss, J.H. (1997) Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses. *Journal of Virology* **71**, 613-623.

Wolff, J.A., Malone, R.W., Williams, P., Chong, W., Acsasi, G., Jani, A., Felgner, P.L. (1990). Direct gene transfer into mouse muscle in vivo. *Science* **247**,1465-1468.

The alphaviruses are a group of about 27 enveloped viruses with a positive sense,

nonsegmented single-stranded RNA genome (Calisher *et al.*, 1980; Strauss and Strauss, 1988). The alphavirus disclosed in this invention, western equine encephalitis virus (WEE), is a member of the WEE antigenic complex and is serologically related to the Sindbis (SIN), Highlands J (HJ), Fort Morgan, Buggy Creek, and Aura viruses (Calisher & Karabatsos, 1988; Calisher *et al.*, 1988). WEE is endemic in western North America and strains/varieties have been isolated from Argentina (AG80-646), Brazil (BeAr 102091) and the former Soviet Union (Y62-33) (Johnson and Peters, 1996; Weaver *et al.*, 1997). In nature, WEE is transmitted from its amplifying hosts or reservoir in wild birds, to man and horses, by mosquitoes (*Culex tarsalis* being the principal vector). While the endemic cycle has resulted in only a limited number of human infections in recent years, in the past, major epidemics of WEE have been recorded. The most extensive epidemic, including 3,336 recognized human cases and 300,000 cases of encephalitis in horses and mules, occurred in the western United States and Canada in 1941 (Reisen & Monath, 1988; Johnson and Peters, 1996).

All alphaviruses share a number of structural, sequence, and functional similarities, including a genome with two polyprotein gene clusters (reviewed in Strauss & Strauss, 1994; Schlesinger & Schlesinger1996). The genomic organization of these viruses is conserved (see Figure 1), with the nonstructural proteins translated directly from the 5' two-thirds of the genomic RNA. A subgenomic positive-stranded RNA (the 26S RNA), is identical to the 3' one-third of the genomic RNA and serves as the translational template for the structural proteins (capsid, E3, E2, 6K and E1).

The nonstructural proteins (nsP1, nsP2, nsP3 and nsP4) are also synthesized as a polyprotein and processed into the four nsPs by a nsP2 protease. Two versions of the nonstructural polyprotein are synthesized in alphavirus-infected cells, due to frequent readthrough of an opal codon between the nsP3 and nsP4 genes in several alphaviruses (Strauss

et al., 1983). The nsPs function in a complex with host factors to replicate the genome and transcribe the subgenomic mRNA. Alphaviruses have characteristic conserved sequences at the extreme 5' and 3' domains and the intergenic region (Ou et al., 1982, 1983; Pfeffer et al., 1998). These conserved domains are required for viral growth and replication and are believed to be important in promotion of protein synthesis and the initiation of RNA-dependent RNA polymerase activity.

The relationship of different WEE isolates to each other has been demonstrated using neutralization tests (Calisher *et al.*, 1988). Additionally, several strains of WEE were typed by oligonucleotide fingerprinting, and found to have greater than 90 % nt homology (Trent & Grant, 1980). The N-terminal sequences of the nucleocapsid, and the E1 and E2 glycoproteins have been determined by Edman degradation, and the E1 and E2 proteins were found to have 82 % and 71 % homology, respectively, to SIN (Bell *et al.*, 1983). Hahn *et al.* (1988) sequenced the 26S region of WEE strain BFS1703. They proposed that WEE originated as a hybrid virus, formed by recombination of an EEE and a Sindbis-like virus, most likely during a co-infection event. They suggested that two crossover events occurred, one within the E3 gene, the other within the 3' nontranslated terminal region (NTR), resulting in a virus whose nonstructural domain, intragenic region, and capsid protein are similar to EEE, with envelope proteins showing homology to SIN.

Weaver *et al.* (1993) sequenced part of the nonstructural domain (nsP2 and nsP3 genes) of strain 5614, demonstrating this area also shows homology to EEE. Short regions within the nsP4 gene and the E1 protein/3' NTR have been determined for many WEE strains, allowing a preliminary assessment of the nucleic acid phylogenetic relationships within the WEE antigenic complex (Weaver *et al.*, 1997). Serological studies (Calisher *et al.*, 1988) and preliminary sequence determination (Cilnis *et al.*, 1996; Weaver *et al.*, 1997) of the HJ genome suggests this

is another closely related virus, and most likely a descendant of the same recombinant viral ancestor as modern WEE.

A highly conserved region of the alphavirus nsP1 gene has been identified, and proved suitable for use in a polymerase chain reaction (PCR)-based genetic assay for alphaviruses, including WEE (Pfeffer et al., 1997). Phylogenetic analysis of this PCR fragment yielded similar results to those obtained by Weaver et al., (1997) for a PCR fragment in the nsP4 gene.

In terms of therapy or prophylaxis, there are very limited possibilities. An inactivated vaccine to WEE is under investigational new drug (IND) status. The vaccine uses formalin-inactivation of cell culture supernatants from WEE-infected tissue culture. It requires a minimum of 3 doses, yearly monitoring of antibody titer and possible boosters. Its effectiveness in the protection against an aerosol challenge of WEE has yet to be established. A WEE live attenuated vaccine based on an infectious clone is under development (J. Smith, personnel communication). The area of DNA immunization is relatively new, and has been reviewed in Hassett and Whitton, 1996; Donnelly *et al*, 1997. Similar to live, attenuated vaccines, DNA vaccines are known to stimulate both humoral and cellular immune responses (Pardoll and Backering, 1997; McCuskie and Davies, 1999). Much of the focus has been on methods to deliver and efficiently express the cloned products. Intramuscular administration of DNA has been one of the original methods used (Wolff *et al*, 1990). A second method uses ballistic delivery of DNA coated gold particles, using high pressure helium gas to propel the particles into the epidermis and dermis of animals (Prayaga *et al*, 1995, reviewed by Robinson *et al*, 1995).

The Applicant identified a number of related areas of research, including the development of subunit vaccines to WEE. In the present invention, the Applicant disclosed the cloning, sequencing and expression of the structural genes of a WEE virus (strain 71V-1658), as described in Netolitzky et al., (2000) "Complete genomic RNA sequence of western equine encephalitis

virus and expression of the structural genes." *Journal of General Virology* **81**, 151-159, which is herein incorporated by reference. The DNA construct (pCXH-3), and a second construct (pVHX-6) were used in DNA immunization studies in a mouse model for protection against intranasal administered WEE.

Summary of the Invention

The present invention is directed to the development of a DNA-subunit vaccine to the WEE virus and its use against such virus. More specifically, DNA to structural components of the WEE virus are expressed and used as the subunit vaccine.

The present invention provides for the complete nucleotide sequence of WEE strain 71V-1658. Two novel cDNA clones, pCXH-3 and pVHX-6 are also disclosed as effective vectors for gene expression.

The present invention also provides the complete nucleotide sequence for the structural gene pcDWXH-7.

It further provides for a process for preparing a recombinant DNA vaccine against WEE virus, comprising cloning and sequencing of 26S region of a WEE virus strain 71V-1658 under conditions suitable to effect *in vitro* transcription and translation of the functional recombinant DNA expression vector pCXH-3 and pVHX-6.

Brief Description of the Drawings

Figure 1. Diagram showing the WEE 71V-1658 sequencing strategy. The location of PCR probe sequences used to screen the WEE cDNA library are also indicated, along with the genomic organization of the virus.

Figure 2. Multiple sequence alignment.

Figure 3. Stem loop structures in the 5' NTR.

Figure 4. Stem loop structures in the 3' NTR.

Figure 5. Phylogenetic relationship of the WEE nonstructural region compared to other alphaviruses.

Figure 6. Expression of WEE structural genes in cell culture.

Figure 7. In vitro transcription and translation of WEE expression vectors.

Figure 8. WEE mouse infectivity model.

Figure 9. Protection using ballistic delivery of pCXH-3.

Figure 10. Protection using ballistic delivery of pVHX-6

Figure 11. Protection using ballistic delivery of pVHX-6.

Detailed Description of the Invention

The complete nucleotide sequence of the 71V-1658 strain of western equine encephalitis (WEE) virus was determined (minus twenty-five nucleotides from the 5' end) and shown in SEQ ID NO: 1. A 5' RACE reaction was used to sequence the 5' terminus from WEE strain CBA87. The deduced WEE genome was 11,508 nucleotides in length, excluding the 5' cap nucleotide and 3' poly(A) tail. The nucleotide composition was 28 % A, 25 % C, 25 % G and 22 % U residues. Comparison with partial WEE sequences of strain 5614 (nsP2-nsP3 of the nonstructural region) and strain BFS1703 (26S structural region) revealed comparatively little variation; a total of 149 nucleotide differences in 8624 bases (1.7 % divergence), of which only

28% of these changes (42 nucleotides) altered the encoded amino acids. Comparison of deduced nsP1 and nsP4 amino acid sequences from WEE with the corresponding proteins from eastern equine encephalitis (EEE) yielded identities of 84.9 % and 83.8 %, respectively. Previously uncharacterized stem loop structures were identified in the nontranslated terminal regions.

A 3100 bp clone was identified (pcDNA-12) from the 3' end of the structural genes. A 1500 bp fragment was PCR amplified and cloned into the 5' end of pcDNA-12 to produce a complete clone of the structural genes (XH-7) as shown in SEQ ID NO: 2. A cDNA clone (pCXH-3) in which the structural genes of WEE strain 71V-1658 were placed under the control of a cytomegalovirus promoter was made, and transfected into tissue culture cells. The viral envelope proteins were functionally expressed in tissue culture, as determined by histochemical staining with monoclonal antibodies which recognize WEE antigens. The construct was used to immunize mice ballistically and intramuscularly. Mice protected ballistically had a significantly reduced risk of infection, against a subsequent intranasal challenge with WEE virus. A new vector was constructed to determine if increased levels of expression could be obtained. The construct used a pVAX vector to express the WEE structural genes (pVHX-6). Upstream portion of the pVHX-6 vector to where it becomes the XH-7 sequence is shown as SEQ ID NO: 3. The remaining nucleotide sequence of pVHX-6 from the point of divergence is identical to that of structural gene pcDWXH-7 of SEQ ID NO: 2.

MATERIALS AND METHODS

Virus Culture and Purification

Tissue culture was maintained in accordance with established methods (Bird & Forrester, 1981). Minimal essential media containing 5 % fetal calf serum (5% DMEM) was used to grow Vero (CRL 1586) and Chinese hamster ovary (CHO) K1 (CCL 61) cells obtained from American

Type Culture Collections. A 10 % suckling mouse brain (SMB) suspension of WEE strain 71V-1658 was kindly provided by Dr. Nick Karabatsos, Centers for Disease Control, Fort Collins, CO. WEE Fleming and California strains were purchased from ATCC (Mannanas, VA). WEE B11 and CBA87 strains were kindly provided by Dr. George Ludwig, United States Army Medical Research Institute of Infectious Disease (Frederick, MD). Seed stocks of WEE strains were made by inoculation of Vero cells with virus suspensions at a multiplicity of infection (MOI) of less than 0.1. For RNA isolation, virus stocks were prepared by infecting Vero cells at a MOI of 10. The virus was precipitated from cleared supernatant by the addition of polyethylene glycol MW 6000 to 7 %(w/v) and NaCl to 2.3 %(w/v). It was subsequently purified on a 20-60 %(w/w) continuous sucrose gradient, followed by resuspension in PBS.

Nucleic Acid Preparation

Viral RNA used in WEE strain 71V-1658 library construction was prepared by the lysis of virus in 0.5 %(w/v) sodium dodecyl sulfate (SDS), and RNA extracted using the cesium chloride/guanidium isothiocyanate method previously described (Sambrook *et al.*, 1989). RNA was precipitated using sodium acetate and ethanol, then stored at -70 °C. Prior to use, RNA was washed with 80 %(v/v) ethanol, dried and dissolved in nuclease-free water (Promega, Madison, WI). Integrity of the RNA was checked on formaldehyde agarose gels (Sambrook *et al.*, 1989). A cDNA library of WEE strain 71V-1658 was made by Invitrogen (San Diego, CA), by the ligation of cDNA into the *Bst*XI site of prepared pcDNAII vector, and electroporation into electrocompetent DH1 F' *Escherichia coli* cells. Manipulation of RNA and DNA followed established procedures (Sambrook *et al.*, 1989; Ausubel *et al.*, 1995). Rapid plasmid preparations were made using the Wizard™ plasmid purification kit (Promega, Madison, WI). Large-scale plasmid preparations used the alkali lysis protocol as modified by Qiagen

(Chatsworth, CA). For PCR, RT-PCR and DNA sequencing, oligonucleotide primer design was guided by information from WEE strain BFS1703 and other partially sequenced WEE strains (Hahn et al., 1988; Weaver et al., 1993), and from regions of sequence conservation (Ou et al., 1982 &1983). Oligonucleotides were synthesized and gel purified either at the Regional DNA Synthesis Laboratory (Calgary, Alberta), or on a Beckman Oligo 1000 DNA synthesizer. A catalog with the sequences of primers used is listed in Table 1.

Construction of pCXH-3

The Invitrogen WEE library was screened by dot blot hybridization (Sambrook *et al.*, 1989) with [^{32}P]-labeled, random primed RT-PCR fragments as probes (Amersham, Oakville, ON). A 3100 bp insert, pcDW-12, was identified, and corresponded to the 3' end of the 26 S RNA. The missing 5' end of the 26S region was generated by RT-PCR using the primers WEE5'Sst1 and WEEP3 (Table 1). The 1500 bp *SstI/NcoI* restricted fragment was inserted into the plasmid, phT3T7BM+ (Boehringer Mannheim, Laval, PQ), to generate a *XbaI* site on the 5' end. The 1500 bp *XbaI/NcoI* fragment was excised, gel purified and inserted into the *XbaI* and *NcoI* restriction sites of pcDW-12. The resulting clone, pcDWXH-7, encoded the complete 26S region of WEE 71V-1658. The structural gene insert from pcDWXH-7 was cloned into the mammalian expression vector, pCI (Promega, Madison, WI). The pcDWXH-7 plasmid was first linearized using *Hind*III, followed by a Klenow fragment reaction to fill in the 5' overhang. The insert was then excised using *XbaI*, gel purified and ligated into the *XbaI/SmaI* digested pCI vector. The isolated recombinant plasmid, pCXH-3, was characterized as having the correct insert by restriction mapping.

Construction of pVHX-6

The clone, pcDWXH-7, encoded the complete 26S region of WEE 71V-1658 was digested with *Sac* I, and religated in the reverse orientation. The isolate, pcDWHX-45, contained the complete 26S of WEE, with the reverse cloning sites (*Hind*III on the 5' end and *Xba*I on the 3' end). The WEE 26S gene segment was excised from pcDWHX-45, and cloned into the *Hind*III and *Xba*I sites of the mammalian expression vector, pVAX (Invitrogen, La Jolla, CA). After transformation into *E. coli* DH10α (Life Sciences, Burlington, ON) and screening of inserts by restriction analysis, a resulting isolate, pVHX-6 was identified. SEQ ID NO: 3 shows the upstream portion of the pVHX-6 vector to where it becomes the XH-7 sequence. The remaining nucleotide sequence of pVHX-6 from the point of divergence is identical to that of structural gene pcDWXH-7 of SEQ ID NO: 2.

Expression of the Structural Genes of WEE

The pCXH-3 expression vector was transfected into Vero or CHO K1 cells using the cationic lipid, Lipofectamine™ (Gibco/BRL, Burlington, ON). Briefly, Vero or CHO K1 cells were grown to 30-50 % confluency in Costar 6-well plates. The monolayers were transfected with pCXH-3 in accordance with the manufacturer's directions, for a period of 5 hrs, followed by a further 29 hr incubation after the addition of 5% DMEM. The monolayers were fixed in methanol:acetone (1:1) for 5 min and washed with PBS containing 0.1 %(v/v) Tween 20 and 3 % BSA (PBS-TB). The cells were incubated 45 min at 37 °C with a 1/100 dilution (in PBS-TB) of concentrated cell supernatant from hybridoma cell lines expressing monoclonal antibodies to the WEE E1 (clone 11D2) or E2 (clone 3F3) proteins, followed by washing with PBS-TB. Monolayers were incubated with a 1/4000 dilution of goat anti-mouse IgG/IgM (H & L) horse radish peroxidase conjugate (Caltag, So. San Francisco, CA) for 45 min at 37 °C. After washing

with PBS-T, 2 mL of TruBlue™ HRP substrate (Kirkegaard & Perry Laboratories, Gaitherburg, MD) was added, and plates were incubated a further 30 min at room temperature, followed by microscopic examination.

In a second method, one-step *in vitro* transcription and translation reactions using the TNT coupled system (Promega Corporation, Madison, WI) was used to express the gene products from both pCXH-3 and pVHX-6, as both have an upstream T7 promoter which can be used for *in vitro* expression of inserts. The RNA was translated in the presence of [^{35}S]methionine to produce radiolabeled WEE proteins, which were further processed with canine pancreatic microsomal membranes. All components of the *in vitro* transcription and translation reactions were incubated together for 90 min at 30 °C. Results were analyzed by SDS-PAGE or radioimmunoprecipitation.

Radioimmunoprecipitation

The TNT reactions were diluted to a volume of 500 ml with RIP buffer consisting of 0.15 M sodium chloride, 0.1% SDS, 50 mM Tris-HCl pH 7.4, and 1% Triton X-100. They were then preabsorbed by incubating with 75 L of protein G-agarose (Gibco BRL) for 30 min at room temperature. The samples were centrifuged at 13,000 rpm for 1 min, and the supernatants were then immunoprecipitated with either 100 µL of supernatants from anti-WEE hybridoma cells or 20 µg of purified anti-WEE antibodies. The reactions were incubated for 1.5 hr at room temperature, after which 75 µL of protein G-agarose was added. The reactions were incubated for an additional 30 min at room temperature. Immunoprecipitated proteins were collected by centrifuging at 13,000 rpm for 1 min. The pellets were washed with 500 µL of RIP buffer and centrifuged at 13,000 rpm for 1 min; this step was repeated three additional times. The pellets were resuspended in 2x Laemmli sample buffer (Bio-Rad Laboratories) containing fresh 2% b-

mercaptoethanol and heated at 100 °C for 10 min. The samples were centrifuged at 13,000 rpm for 1 min, and the supernatants were collected. The immunoprecipitated [^{35}S]labeled WEE proteins were further analyzed by SDS-PAGE and autoradiography. Radiolabelled [^{14}C]molecular weight markers from Amersham Pharmacia Biotech were also run on the polyacrylamide gels.

DNA Sequencing

Automated sequencing of the 26S region was performed using the ABI Prism Dye Terminator Cycle Sequencing or Big-Dye™ Terminator Cycle Sequencing kits of plasmid templates according to the manufacturer's instructions (PE-Applied Biosystems, Foster City, CA). Sequencing reactions were purified on Centri-Sep™ columns (Princeton Separations, Adelphia, NJ), dried and analyzed on an ABI 373 or 310 automated sequencer. For the nonstructural region, template cDNAs were generated in a single-step integrated RT-PCR procedure using the Titan™ RT-PCR kit (Boehringer Mannheim, Laval, PQ), following the manufacturer's suggested protocols. RT-PCR products were purified using the QIAquick™ PCR Purification kit (Qiagen, Chatsworth, CA) and sequenced (50-100 ng DNA per reaction). The extreme 5' end of the genome was not sequenced in WEE 71V-1658. However, a 5' RACE reaction (Frohman et al., 1988) was used to obtain a cDNA fragment from the 5' terminus of WEE strain CBA87. Briefly, primer WEE559 (GGTAGATTGATGTCGGTGCATGG) was used to prime reverse transcription of the 5' terminus of the viral RNA. After poly(A) tailing of the cDNA with terminal transferase, a plus sense primer (GTACTTGACTGACTGTTTTTTTTTTTTTT) was used in conjunction with WEE559 for amplification of the 5' terminus.

Nucleotide Sequence Analysis and Assembly

Sequence traces were edited manually and assembled using the Seqman component of the Lasergene DNA analysis software (DNASTAR, Madison, WI). Codon preferences and patterns were assessed using the CodonUse and CodonFrequency programs, while the overall frequency of mononucleotide and dinucleotides was calculated using the Composition program (Wisconsin Package, Version 9.0, Genetics Computer Group, Madison, WI). Quantitative assessments of sequence similarities (nucleotide and amino acid), were calculated by preliminary alignment using the Pileup program, followed by manual alignment adjustment, and analysis with the Distances program (GCG). Amino acid sequences aligned as described, were used as the basis for generating phylogenetic trees (GCG). The GeneQuest module of the Lasergene program (DNASTAR, Madison, WI) was used to predict and calculate RNA secondary structures at the ends of the genomic RNA using minimal energy calculations. Multiple sequence alignments were accomplished using the Clustal component of MegAlign (DNASTAR). The complete WEE genomic nucleotide sequence has been submitted to GenBank (Accession Number AF143811).

Administration of DNA or Inactivated Virus

DNA solutions or an inactivated WEE virus vaccine in PBS, were administered to the mice by ballistic or intramuscular (IM) routes. For IM route of administration, a 27 g needle was used to deliver 50 μg of DNA (pCXH-3 or pCI - negative control) or 50 μL of inactivated WEE vaccine (SALK WEE inactivated vaccine). The volume of inoculum used was 100 μL, diluted in PBS. Fifty μL was administered IM to each of the hind leg muscles of a mouse. When boosters were given, they were administered 14-28 days apart. For ballistic administration, mice were shaved in the abdominal area with electric hair clippers. The mouse was subjected to

ballistic delivery of DNA coated onto gold particles following the manufacturer's standard specifications. The Helios Gene Gun (Biorad, Mississauga, ON) was used as directed, at a pressure setting of 400 psi. Mice were given 1.25 μg DNA and 0.5 mg gold, 1 μm diameter, per shot, and up to three shots for one dose time. Boosters were given 14-28 days apart. The mice were challenged 14-28 days after the final booster.

Mouse Infectivity with WEE

Female BALB/c mice, 17-25 g, were obtained from the mouse breeding colony at Defence Research Establishment Suffield (DRES), with the original breeding pairs purchased from Charles River Canada (St. Constant, Quebec, Canada). The use of these animals was reviewed and approved by Animal Care Committee at DRES. Care and handling of the mice followed guidelines set out by the Canadian Council on Animal Care. Virus was administered to the mice by intranasal (IN) or intraperitoneal (IP) routes. The volumes of inoculum used were 50 μL for IN and 100 μL for IP. For IN administration, mice were anaesthetized with sodium pentobarbital (50 mg/kg body weight, intraperitoneal). When the animals were unconscious, they were carefully supported by hands with their nose up, and the virus suspension in PBS was gently applied with a micropipette into the nostrils. The applied volume was naturally inhaled into the lungs. For IP infection, the mouse was manually restrained, and a 1 ml tuberculin syringe fitted with a 27 g needle was used to administer approximately 100 μL of the virus suspension in PBS. Infected animals were observed daily, for up to 14 days post infection.

RESULTS

Complete Nucleotide Sequence of WEE Genome and Deduced Amino Acids

The nucleotide sequence of WEE strain 71V-1658 (SEQ ID NO: 1) was determined via several distinct sequencing strategies, as summarized in Figure 1. The 5' terminus of 25 nt was not determined for this strain. However, it was determined by sequencing a 5' RACE product from strain CBA87. Excluding the terminal 5' cap structure and the 3' poly(A) tail, the genomic sequence of WEE was found to be 11,508 bases long. The base composition was 28 % A, 25 % C, 25 % G, and 22 % U. The dinucleotide usage of the WEE genome was compared with those values anticipated from the base composition. Several dinucleotides were found in lower proportions than anticipated, notably UpA (81%), CpG (83%) and CpC (85%) (data not shown). Codons containing the CpG dinucleotide were present at 82% of the anticipated value, including codons for serine (78%), proline (80%) and arginine (78%).

The WEE 71V-1658 sequence was used to conduct a variety of phylogenetic analyses with previously determined alphavirus sequences. The alphaviruses used in the analyses included EEE strain North American variant (Genbank Acc. No. X67111), O'Nyong Nyong (ONN) strain Gulu (Genbank Acc. No. M33999), Ross River (RR) strain NB5092 (Genbank Acc. No. M20162), Semliki Forest (SFV) (Genbank Acc. No. J02361), SIN strain HR (Genbank Acc. No. J02363) and VEE ID (Genbank Acc. No. L04653). The degree of conservation among the various sequences (nucleotide and amino acid) through the stereotypical alphavirus genome is shown in Table 2. The carboxy-terminal domain of nsP3, which consistently fails to exhibit homology among sequenced alphaviruses, was excluded from this comparison as it has been adjusted for in previous analysis (Weaver et al., 1993). The deduced amino acid sequences for nsP1-4 of WEE 71V-1658 demonstrated closest identity to the corresponding proteins from EEE (Table 1), reflecting similar observations made for nsP2 and nsP3 of WEE 5614 and EEE (Weaver et al., 1993).

Nontranslated Terminal Regions

Alignment of the 5' terminal nucleotide sequences of WEE CBA87 and WEE 71V-1658 is shown in Figure 2a, along with a comparison of the 5' termini from EEE and VEE. The close similarity between WEE and EEE, has been verified experimentally, in that a EEE/Highlands J degenerate primer, EHJ5', was able to PCR amplify the 5' end of the WEE genome, while an analogous SIN primer could not (data not shown).

Potential stem loop structures were found in WEE 71V-1658, including a stem loop at the extreme 5' terminus (2-30) and a pair of stem loops (137-189) (Figure 3a). The homologous structures for EEE are also shown (Figure 3b) (Ou et al., 1983). Minimal energy values calculated for the stem loops were similar between WEE and EEE. Further analysis of the region between the structures described above, indicated a large, highly base-paired stem loop structure (39-131), that had not been previously described, and was observed in SIN and EEE in a similar location (data not shown).

The sequence of WEE 71V-1658 3' NTR, overall, shared little homology with any of the alphaviruses examined, but included the highly conserved 19 nt region at the 3' end (11490-11508), which was identical to that determined for WEE BFS1703 by Hahn et al., 1988. Two copies of the characteristic 40 base Sindbis-like terminal repeats as previously reported (Ou et al., 1982) were found in WEE 71V-1658 (11234-11273 and 11292-11331). However, the 3' NTR of WEE showed some surprising results that had not been previously described. The first 40 nt terminal repeat formed the backbone for the formation of a 57 nt double stem loop structure (11228-11284) (Figure 4b), consisting of an α and β loop. The second 40 nt repeat of WEE formed a nearly identical 59 nt double stem loop structure (11285-11343), directly adjacent to the first structure. SIN with three 40 nt repeats, forms three double stem loops (Figure 4a) while EEE, which does not contain a SIN-like 40 nt repeat, contains the α and β loops (Figure 4c).

Nonstructural Region

Comparisons within the nonstructural regions (4475 nt) of WEE strains 71V-1658 and 5614 (Weaver et al., 1993), yielded 94 nt changes resulting in 26 amino acid substitutions (1.8% difference) as summarized in Table 2. The most notable variation, a three-base deletion (4530) within the nsP3 gene of WEE 71V-1658 constitutes the only insertion/deletion observed within the polypeptide encoding regions. A short hypervariable region was observed (1421-1449), where 11 of 28 nt were different between the two WEE strains (Figure 2b). The presence of an opal termination codon and partial read-through site at the junction of nsP3 and nsP4 is consistent with WEE 5614. Extending previous phylogenetic analyses of WEE (Weaver et al., 1993, 1997), phylogenetic trees depicting viral relatedness were constructed with the Distances program (GCG), for the unexamined genes (nsP1, nsP4) and the entire nonstructural polypeptide encoding region (Figure 5). The data reveals the close relationship of WEE to EEE, relative to the other alphaviruses analyzed.

Structural Genes

The largest WEE cDNA clone isolated, pcDW-12, was 3100 bp in size, but missing 5 nt and the poly(A) tract from the 3' end as determined by restriction mapping and DNS sequence analysis. The missing 5' 1500 bp fragment was synthesized using PCR (primers WEE5'Sst1 and WEEP3) and subsequently cloned into pcDW-12 to yield a full-length clone of the structural genes (pcDWXH-7) (SEQ ID NO: 2). Comparison of the structural region of WEE 71V-1658 with WEE BFS1703 (Hahn *et al.*, 1988), indicated 53 nt changes, resulting in only 11 amino acid differences, of which two were nonconserved. One difference in residue was observed from the amino acid sequence of the N-terminus of the E2 protein of the WEE MacMillan strain (Bell *et*

al., 1983), when this was compared to the deduced protein sequence of 71V-1658. A short fragment (802 nucleotides) of the WEE 71V-1658 E1 protein gene, and the 3' NTR had been published previously (Weaver *et al.*, 1997); comparison with the sequence reported herein indicated no differences.

Expression of Structural Gene

Expression of the insert from the cytomegalovirus (CMV) promoter was accomplished by transfection of the pCXH-3 plasmid into either Vero or CHO K1 cells. Cells expressing the E1 or E2 proteins were detected through the use of specific E1 or E2 monoclonal antibodies to WEE, followed by histochemical staining with the HRP substrate, Tru-Blue as demonstrated in Figure 6a. The control cells transfected with pCI alone showed no staining (Figure 6b), thus, demonstrating the fidelity of the proteins translated from the cloned 26S region. *In vitro* translation of the insert using TNT T7 rabbit reticulysate and canine microsome system demonstrated synthesis of ^{35}S-methionine-labelled proteins of the correct size as indicated by immunoprecipitation with monoclonal antibodies to the NC, E1 and E2 proteins (data not shown). Similarly, the construct pVHX-6 was along demonstrated to produce the correct MW proteins as determined by *in vitro* transcription/translation. The level of expression for pVHX-6 was significantly higher then for pCXH-3 (Figure 7).

Protection Against WEE Infection Using DNA Immunization

Different strains of WEE were shown vary in their virulence in BALB/c mice. When similar amounts of WEE were given intranasally to BALB/c mice, time to death varied from 4 to 8 days. The California and Fleming strains were the most virulent (Figure 8), and the Fleming strain was chosen as the challenge strain in protection studies. IP administration of the virus did

not kill adult mice (data not shown). Intramuscular administration of pCXH-3 did not show any protection, using one or two doses of 50 µg, followed by challenge 30 to 90 days after the final dose (data not shown). Intramuscular administration did result in an increase in antibody titre to WEE as determined by ELISA using a monoclonal antibody to the E1 protein of WEE (data not shown). Expression and protection of pCXH-3 DNA when delivered ballistically. pCI was used as a control DNA. When two doses of pCXH-3 was given, protection of 50% was demonstrated as compared to no protection for pCI (Figure 9) or PBS controls (data not shown). IM injection showed marginal protection (one group 25% survival - data not shown). The dose of WEE Fleming strain (challenge strain) was 1.25 x 10^4 PFU for 100% killing via an intranasal route of infection. Preliminary studies examining protection using the pVHX-6 vector, indicated promise with this construct using the Gene Gun, and ballistic delivery. With the pVHX-6 vector, one mouse succumbed immediately to the effects of the sodium pentabarbital (anaesthetic). The remaining three mice showed no signs of coming down with a WEE infection, and remained completely heathy (Figure 10). Of the four pVAX control mice, all showed signs on WEE infection, and two of the four mice died, while two did recover. A repeat of this experiment using 3 or 4 doses of pVHX-6, given 2 weeks apart, showed complete protection of the mice, similar to 3 doses of WEE inactivated vaccine (Figure 11). Three or 4 doses of pVAX showed results similar to the saline control, with only about 60% of the mice surviving Figure 11.

DISCUSSION

The WEE 71V-1658 genomic sequence of 11,508 bases was determined directly from cDNA clones of WEE or via sequencing RT-PCR products. The first 25 bases of the WEE genome was determined indirectly, through the use of a 5' RACE reaction in WEE CBA87. Noting the relatively high conservation in the WEE sequences overall (1.7% divergence) and in

the overlap region between the two WEE sequences (see Figure 2a), it appears that the 5' ends of 71V-1658 and CBA87 are of similar size and sequence.

Comparison of WEE 71V-1658 to other partial sequences of WEE (Hahn et al., 1988; Weaver et al., 1993) suggests little variation at the nucleotide level among these viruses (Table 2), showing an overall nt sequence difference of 1.7 % over 8624 nt. Given a calculated rate of divergence of 0.028 % per year for the WEE E1 protein (Weaver et al., 1997), the expected nt divergence for a difference in isolation of 18 years between the strains, should be 0.5 % (71V-1658 isolated in 1971 and BFS1703 in 1953). The E1 protein itself showed a rate of divergence of 1.5% in nt sequence between 71V-1658 and BFS1703. The lower rate observed by Weaver et al., (1997) could be due to greater conservation of structure at the C terminus of E1, from where the rates of divergence were calculated. Areas with high rates of divergence were observed between WEE strains 71V-1658 and 5614 at the 3' end of nsP1 and the 5'end of nsP4 (Table 2). The relatively high interstrain value for nsP1 (4.5% difference) may be due to the presence of a small hypervariable region, with 11 of 28 nt changed in strain 5614 (Figure 2b). Variation in nsP4 occurred in a stretch of 21 nt at the 3' end of the 5614 sequence, and were left out of subsequent homology comparisons (similarity with the EEE sequence was maintained in this region). Discounting the carboxy-terminal region of nsP3 also gives a more accurate picture of the homology of the nsP1-4 nonstructural region (Weaver et al., 1993). The results for comparison of nt and protein sequences of WEE to other alphaviruses is shown in Table 2, and are similar to those obtained with nsP2 and nsP3 of 5614, when compared to other alphavirus sequences. Phylogenetic analysis of the WEE 71V-1658 deduced protein sequences of nsP1, nsP4 and the nsP1-4 region, as related to other alphaviruses (Figure 5), illustrates the close relationship to EEE (HJ sequences were very limited for comparative purposes and were not included).

Assessments of codon usage frequencies and the frequency at which certain dinucleotides are found throughout the genome identified a number of statistical anomalies. The slight CpG dinucleotide deficiency previously described within other alphaviruses, and WEE itself, was confirmed in this study, at levels comparable to those reported (Weaver et al., 1993). The CpG under representation is a typical feature of vertebrate genomes, and is not seen in invertebrates. Viruses which infect dual hosts, such as the arboviruses, might be expected to utilize an intermediate nucleotide bias, as indicated by the slight CpG under-utilization observed in alphaviruses (Weaver et al., 1993). A pronounced under-representation of two other dinucleotides was also observed within the WEE genome, UpA, and CpC, a phenomenon noted throughout the genome, though the role of these codon preferences is unclear.

The 5' NTR sequence of WEE shows a close phylogenetic affiliation to EEE, and to HJ, although the HJ sequence information is more limited. Ou et al., (1983) had previously predicted (based on minimal free energy calculations) two hairpin structures at the 5' NTR of several alphaviruses including SIN and EEE. Both structures are present in WEE, the first of which is a 5' terminal hairpin structure (2-30), similar to that calculated for EEE (Figures 3a and b). The second is a dual hairpin structure (137-162, 165-189) which is almost identical to that identified for EEE. The region between the terminal and dual hairpins can itself form a long hairpin structure, and includes highly conserved stretches of 92 nt (data not shown). The significance of these structures is currently unknown.

Previous reports (Hahn et al., 1988; Pfeffer et al., 1998) suggested WEE virus arose as a result of two recombination events between alphavirus-like ancestral viruses. The first recombination occurred near the junction of the E3 and capsid genes. The second recombination occurred 80 nucleotides from the 3' end of the genome. Evidence for the occurrence of the second recombination event is inferred from sequence similarities of the 3' NTR between WEE,

EEE and SIN, in which WEE shows greater similarity to EEE (65 %) than to SIN (50 %) in the last 100 nt of the 3' end. However, the apparent plasticity of the 3' NTR may only be reflecting the selective pressures under which the nascent WEE virus evolved, resulting in rapid selection of 3' sequences which are more similar to EEE, and may not represent an actual recombination event as previously postulated.

The 3' NTRs of alphaviruses are characterized by widespread sequence divergence and yet contain small, strongly conserved motifs (reviewed in Strauss & Strauss, 1994; Pfeffer et al., 1998). Analysis of the 3' NTR indicated the presence of double stem loop structures among SIN and WEE (Figures 4a and b). Interestingly, the 40 bp repeat found in SIN and WEE is contained within the double stem loop structure. SIN was found to contain 3 double stem loop structures and WEE was found to contain two. In SIN, the spacing between the three double stem loop structures was around 30 nucleotides, while in WEE the distance was zero nt separating the structures. Additional alphaviruses were assessed and it is interesting to note that double stem loop structures were found in many of the WEE- and SIN-related viruses (SIN, Aura, Babanki, Ockelbo, Kyzylagach, Whataroa, WEE and HJ). The double stem loop structures found in SIN and WEE viruses consisted of the α loop (AUGUA[U/C]UU) and the β loop (GCAUAAU) (Figure 4b). Surprisingly, while EEE does not have the 40 bp repeat element found in SIN and WEE, it contains the α and β loop structures (Figure 4c). The significance of these conserved loop structures between SIN, WEE and EEE viruses has yet to be elucidated, although previous studies suggest a role in viral replication and/or host specificity (Kuhn et al., 1990; Kuhn et al., 1991). For example a deletion of 26-318 nt from 3' end of SIN, resulted in reduced viral replication in mosquito cells but not in chicken cells (Kuhn et al., 1990). In contrast, substitution of the SIN 3' NTR with the substantially different RR 3' NTR (which lacks the 40 bp repeat and double stem loop structures), had no effect on the growth of the chimeric virus in mosquito cells,

suggesting that host proteins interact with the 3' NTRs to cause differential host effects (Kuhn et al., 1991).

The 26S region of 71V-1658 was placed under the control of the CMV promoter of pCI. To test for functional expression of the pCXH-3 vector and for a functional product in cell culture, the pCXH-3 vector was transiently transfected into Vero cells. WEE proteins were detected on the cell using specific monoclonal antibodies to both the E1 (Figure 6a) and E2 proteins (data not shown). The binding specificity of these monoclonals has been previously determined by western blot analysis and immunoprecipitation analysis (data not shown). The use of pCXH-3 in DNA immunization experiments indicated that the construct could partially protect against WEE intranasal challenge using ballistic delivery. Preliminary results do indicate that WEE reactive antibodies can be detected by ELISA when the pCXH-3 plasmid is given intramuscularly (unpublished results). However, this afforded no protection to the mice, as there were no survivors. Intranasal (data not shown) delivery of the pCXH-3, with and without liposome encapsulation did not demonstrate any protection under the conditions used. Mice immunized with the pCI control plasmid did not show any signs of protection in these studies.

Expression of the WEE structural proteins in the pCI-based vector (pCHX-3) gave moderate to poor levels of expression in vitro, using the TNT expression kit. A new vector, pVAX (Invitrogen) was designed for DNA immunization and was basically the same as pCI, but lacked the intron found in the pCI vector. Initial restriction mapping of pCXH-3 indicated the plasmid was the expected size, but later analysis indicated a extra 4 kb fragment was present (data not shown). The WEE structural proteins were cloned and expressed in pVHX-6, indicating the correct sized proteins by SDS-PAGE, and producing higher levels of WEE product *in vitro* (Figure 7). Preliminary results with pVHX-6 indicated it could completely protect mice against an intranasal challenge of WEE. While 50% of the pVAX mice did survive, they all

demonstrated at least moderate to severe infection with WEE. It is possible that pVAX contains CpG motifs that show some protective effect, through a nonspecific adjuvant like effect (Kreig *et al*, 1998). However, there was a dramatic difference between the pVAX and the pVHX-6 group, in the protection afforded the two groups of mice.

The plasmids, pCXH-3 and pVHX-6 show promise as vaccine candidates for WEE. This is especially important for protection against an aerosol challenge of WEE, and event that would be envisioned in a potential biological warfare attack using WEE as a biological warfare agent. This agent is difficult to protect against if delivered aerosolly, as the agent is purported to travel up the nerves directly into the brain. The research is applicable to VEE and EEE, as these viruses can also cause encephalitis following a similar route of infection (equines and potentially human).

It is to be understood that the embodiments and variations shown and described herein are merely illustrative of the principles of this invention and that various modifications may be implemented by those skilled in the art without departing from the scope and spirit of the invention.

Table 1. WEE 26S Region Primers

Name	Length	Sequence
WEEPRO	30	AATCACCCTCTACGGCTGACCTAAATAGGT
WEEPR-SST	24	GGCTGAGCTCAATAGGTGACGTAG
WEE3'	30	GTAGTGTATATTAGAGACCCATAGTGAGTC
WEE5'SST	20	TCCAGATACGAGCTCATACT
WEEN1	20	GGTGCCGCTGGAGGCCGTTT
WEEN1A	20	GATCTTAGGAGGTCGATAGC
WEEN2	20	GGCTGATGAAACCACTCCAC
WEEN3	20	CCACCCGTGTGCTATTCACT
WEEN3A	20	CGCCGTGTTTCAGCCCAATA
WEEN4	20	TCACGAGCGGAGCATCTGAG
WEEN5	20	GGCATCACCCTCCACCTGAC
WEEN6	20	TTGTTATTCTGTTCCGCTGC
WEEN7	20	CTATTGATCATGCAGTCGCA
WEEN8	20	AGTGGAGCCTCTGCGAGCGT
WEEN9	20	GAGGAGTGGGCGGGAAAGGC
WEEN10	20	CTAAAACTCGATGTATTTCC
WEEN11	20	ACGCGAACGAAGATGAACGG
WEEN12	20	ACTGTCATTGTGCTGTGTGG
WEEN13	20	CACAGTCATTCCTTCACCAC
WEEN14	20	CGTCATCAGAAAGGGGCTTG
WEEN15	20	CAAAGCTGACAGGGAGGGAC
WEEN16	20	GGAAAGCTGGTAAAGTGCCA
WEEN0	20	GGAGAACCACATAAAGTCGA
WNSP1	25	GGCTAACGTGGACAGGGACGTGATG
WEEP0	20	GGCTATCGACCTCCTAAGAT
WEEP0A	20	CTGTCGGTTCCCTGGTTTAG
WEEP1	20	CTGGGGAACGTCGCCATACT
WEEP2	20	CGTTCTCCAGCAGCGTGTCG
WEEP2A	20	TATTGGGCTGAAACACGGCG
WEEP3	20	CTTCAAGTGATCGTAAACGT
WEEP4	20	ACTCCAGCCCTTCTCGCCCC
WEEP5	20	GTTCGACCAACGCCTTATAC
WEEP6	20	AAGGGTGAAAAAGCGGCTGA
WEEP7	20	GGTGATTCTGATGATCTCAC
WEEP8	20	TGGAAACTGCCGCCTGGAAT
WEEP10	20	CCTTGATGTCATGGTCGTGG
WEEP11	20	TGCACTGAGTGGTCTGTGTG
WEEP12	20	ATGTTTCAGCGTTGGTTGGC
WEEP13	20	GTGTTCTCACTGTCACAGAA
WEEP14	20	ATGTGTGGTCGCTTCCTTCA

Table 2. Percentage Variation in Nucleotide and Encoded Amino Acid Sequences Between WEE 71V-1658 and Other Alphaviruses

	WEE (BFS1703)	WEE (5614)	EEE	VEE	SIN	RR	ONN	SF
5′ NTR	-	-						
nsP1 (nt)	-	(4.5)	25.1	34.8	40.9	37.8	39.7	39.1
nsP1 (aa)	-	(6.3)	15.1	32.1	40.3	35.5	37.2	33.3
nsP2 (nt)	-	1.8	28.2	34.6	43.9	42.1	42.9	42.8
nsP2 (aa)	-	0.6	16.2	26.5	44.9	43.2	44.9	44.4
nsP3 (nt)*	-	1.8	30.2	36.7	45.8	39.3	42.6	42.2
nsP3 (aa)*	-	2.1	18.8	32.4	46.3	38.7	40.9	43.5
nsP4 (nt)	(1.8)	(2.4)	25.6	31.4	34.7	35.3	36.0	37.0
nsP4 (aa)	(2.6)	(4.3)	11.7	21.4	26.8	27.3	25.8	27.4
intervening (nt)	4.3	-	56.6	51.5	47.6	44.7	60.0	47.7
Capsid (nt)	2.1	-	26.3	40.8	47.7	46.3	47.5	48.2
Capsid (aa)	1.5	-	16.8	43.5	52.8	53.3	54.6	54.3
E3 (nt)	1.1	-	45.6	40.7	38.3	51.7	47.5	46.7
E3 (aa)	1.7	-	38.0	39.6	39.4	46.0	45.8	43.9
E2 (nt)	1.2	-	51.2	52.3	36.2	51.7	55.3	52.8
E2 (aa)	1.0	-	59.0	60.0	31.7	63.5	65.7	64.7
6K (nt)	0.6	-	53.3	46.3	26.1	51.9	50.3	54.3
6K (aa)	1.8	-	65.6	59.3	32.7	72.2	69.1	75.9
E1 (nt)	1.5	-	43.8	45.8	29.6	47.2	48.5	44.4
E1 (aa)	0.5	-	49.0	51.0	23.4	51.5	54.8	50.3
3′ NTR (nt)	0.7	-	57.8	55.0	53.2	69.1	65.8	60.3

* based on N terminal domain, C terminal domain discarded due to lack of homology between alphaviruses

( ) based on incomplete sequence data: nsP1 (289 nt) and nsP4 (207 nt for BFS1703, 113 nt for 5614)

- no data